Towards a capabilities database to inform inclusive design: experimental investigation of effective survey-based predictors of human-product interaction

Raji Tenneti [a] [*], Daniel Johnson [b], Liz Goldenberg [c], Richard A. Parker [d], Felicia A. Huppert [e]

[a] *Department of Psychiatry, University of Cambridge, Cambridge CB2 0QQ, UK*
[b] *Faculty of Science and Technology, Queensland University of Technology, Brisbane 4001, Australia*
[c] *School of Psychology, University of New South Wales, Sydney 2052, Australia*
[d] *Department of Public Health and Primary Care, University of Cambridge, Cambridge CB2 0SR, UK*
[e] *Department of Psychiatry, University of Cambridge, Cambridge CB2 0QQ, UK*

A key issue in the field of inclusive design is the ability to provide designers with an understanding of people's range of capabilities. Since it is not feasible to assess product interactions with a large sample, this paper assesses a range of proxy measures of design-relevant capabilities. It describes a study that was conducted to identify which measures provide the best prediction of people's abilities to use a range of products. A detailed investigation with 100 respondents aged 50-80 years was undertaken to examine how they manage typical household products. Predictor variables included self-report and performance measures across a variety of capabilities (vision, hearing, dexterity and cognitive function), component activities used in product interactions (e.g. using a remote control, touch screen) and psychological characteristics (e.g. self-efficacy, confidence with using electronic devices). Results showed, as expected, a higher prevalence of visual, hearing, dexterity, cognitive and product interaction difficulties in the 65-80 age group. Regression analyses showed that, in addition to age, performance measures of vision (acuity, contrast sensitivity) and hearing (hearing threshold) and self-report and performance measures of component activities are strong predictors of successful product interactions. These findings will guide the choice of measures to be used in a subsequent national survey of design-relevant capabilities, which will lead to the creation of a capability database. This will be converted into a tool for designers to understand the implications of their design decisions, so that they can design products in a more inclusive way.

Keywords: inclusive design, design practice, design tools, user-centred design, human factors

1. Introduction

Capabilities are fundamental attributes that a person needs to use everyday products. When interacting with a product, demands will typically be made on sensory (such as vision, hearing), motor (such as dexterity, locomotion, reach and stretch) and cognitive (such as memory, learning, comprehension) capabilities. In the context of design, capability refers to an individual's level of functioning, from very high ability to extreme impairment, which has implications for the extent to which they can interact with products (Johnson et al, 2009). As the human body ages, especially beyond the age of 65 years, there is a substantial reduction in functional capability (motor, sensory and cognitive capabilities) (Huppert, 2003). Age-related decline has implications for design. Failure to take account of this reduced functional capability in the design process results in older people, who constitute a growing proportion of the adult population, becoming excluded from product use (Elton and Nicolle, 2010). Inclusive design is a design philosophy that aims to consider this reduced functional capability during the design process, with the aim of making products functionally accessible to and usable by as many people as reasonably possible. By meeting the needs of those who are often excluded from product use, inclusive design improves product experience across a broad range of users (Coleman, 2001).

One way of promoting a better understanding of user needs is through the provision of end-user data, such as anthropometrics (e.g. physical characteristics) and capabilities databases

[*] Corresponding author. Tel: +44 1223768507; Fax: 1223 336968; Email: raji.tenneti@gmail.com

for design of environments and products (McGinley et al, 2010). It should be noted that an end-user database in itself is unlikely to be of use to designers and that tools need to be developed that present the relevant data in an accessible and useful way for predicting difficulty and exclusion from product use. Examples of such tools include ADAPS (Molenbroek,1987), a computer-aided design model which uses twenty-five functional body dimensions of 822 elderly Dutch people, HADRIAN (Porter *et al.*, 2004), a computer-aided design tool which allows evaluation of products and services against a database which uses 3D anthropometry and functional abilities and the Exclusion Calculator (Clarkson et al, 2007), a tool designed to estimate the number of people who would be excluded from using a particular product, based on assessing the demands on each individual capability domain.

Tools for predicting difficulty and exclusion need to be able to give designers a picture of the full range of capabilities and also the ability to consider and understand the multi-dimensional nature of capability profiles (Johnson et al, 2009). For example, it may be important to know not only how many people will have difficulty with the vision or hearing demands of a product, but also how many people will have difficulty with neither or with both. To obtain such information requires extensive measurement of people's capability across a range of domains (e.g. vision, hearing, dexterity, reach and stretch, locomotion, communication, thinking). The best way to measure these capabilities for the prediction of difficulty with products is not yet known. The breadth and multi-dimensional nature of capabilities can be best captured and represented through a database that covers multiple capability domains for a representative sample of the population.

1.1 Limitations with existing databases

A number of problems exist with the currently available end-user capability databases that have implications for their value in estimating the capabilities of the population. Consideration of these issues is instrumental in identifying the key features of future surveys designed to create a reliable capability database to inform the measurement of inclusion in product designs. Some of the problems associated with these databases, identified by Johnson et al (2009) include:

- Lack of data on multiple capabilities. Existing databases such as Adultdata (Peebles and Norris, 1998), Older Adultdata (Smith et al, 2000) and Childata (Norris and Wilson, 1995), which cover multiple domains in a single publication, draw their data for each capability domain from different samples and thus assessment of multiple capabilities is not possible.
- Absence of surveys with an appropriate level of specificity in the questions. Where existing health and disability surveys are used, they ask only general questions and disease-specific questions, which are not very useful to Inclusive Design, since knowing that someone suffers from a particular disease (e.g. diabetes) does not reliably provide an indication of their capabilities. Surveys that fall into this category include the General Household Survey, the Family and Children Study, the Family Resources Survey, the Labour Force Survey, the Omnibus Survey and the Census (Bajekal et al, 2004, European Commission, 2008).
- Data derived from a non-representative sample of the population. For example, Geron 1998 Dutch Elderly study was biased to high educational level of the sample (Steenbekkers and van Beijsterveldt, 1998) and the ONS (Office for National Statistics) Great Britain Disability Follow-up Survey 1996/97 (Grundy et al, 1999) is limited by problems with the sift criteria used to sample the population (e.g. certain age brackets are known to be under-represented). In addition, the Disability Follow-up Survey was

designed to provide a measure of severity of disability and not intended for use in providing a full range of capability estimates across the normal population.

1.2 Design-relevant survey of capabilities

In light of the above-mentioned limitations with the existing surveys, it is evident that a design-relevant survey of capabilities is needed in order to build a capability database. The key question is what measures can be devised that provide the most accurate and generalisable predictors of difficulty or exclusion when interacting with products. Johnson et al (2009) reviewed the potential influences on the measures of capability and concluded that a number of issues need to be considered for the construction of a survey to reliably assess capabilities. Specific issues include: self-report versus performance measures; granularity of measurement; psychological characteristics; and naturalistic versus experimental settings for performance. These are discussed below. For more information on these issues, please see (Johnson et al, 2009).

Self-report versus performance measures - A person's capability can be assessed through either their own reports of capabilities or product interactions or objective measures of their performance. While self-report measures rely on the accuracy of the respondent's judgments and are easier to administer and less expensive, objective performance measures of capabilities require specialised equipment and can be time consuming (Kivinen et al, 1998; Hupkens et al, 1999). The two types of measures potentially assess different aspects of capability, so it is informative to know how well each of them predicts people's experienced difficulty or exclusion when interacting with products.

Granularity of measurement - Whether self-report or performance measures are used, the granularity of the measurement needs to be considered. For example at the lowest level of granularity, measures could be taken of a component function (such as vision measured by an eyesight test); at a medium level of granularity, measures could be taken of a specific activity (such as reading the LCD on a mobile phone); or at a higher level of granularity, measures could be taken of a task which integrates number of functions and activities. In line with Johnson et al (2009), we use the term component function to refer to the basic sensory, motor and cognitive capabilities, which provide data on an individual's capabilities independently of how these capabilities are used to interact with products. Component activities are defined as smaller tasks within the larger product interaction. All these types of measures are arguably relevant to designers. Therefore, it is important to know the extent to which each of these provide a good prediction of how well people interact with products.

Psychological characteristics - Certain psychological states and traits of a person are likely to directly influence their capability when using a product. There is extensive evidence that mood or emotional state can change a person's perception, thoughts and behaviour (see Forgas, 2008 for a review). Similarly it can be expected that greater self-confidence and motivation will increase a person's capability. The success of product interactions can be influenced by the beliefs and attitudes that a person holds. These include: self-efficacy, self-esteem, optimism and perceived mastery. However, there is a distinction between general self-efficacy and product-related self-efficacy. A question exploring general self-efficacy would be, 'I can solve most problems if I invest the necessary effort', whereas a question exploring product-related self-efficacy would be, 'I am confident in my current skills and ability to learn how to use a new piece of equipment in my home'. In other words, an individual may have high general self-efficacy but poor self-efficacy regarding technology use. Therefore there is a need to establish which is the better or best predictor of success when using a product.

Performance: naturalistic versus experimental settings - A person's ability to perform an activity can be shown to vary between the home and a laboratory (Kivinen et al, 1998). In the home, people may develop coping strategies to enable them to perform certain tasks despite capacity limitations. Conversely, in an experimental setting, people may push themselves during a performance test in a way that they would not be able to do on a regular basis. Therefore, it is important to know whether performance measured in the home or in a controlled testing environment provides the better prediction of people's capability when interacting with products.

This paper describes research conducted in the form of an exploratory study that aims to provide initial answers to the issues identified above. The research represents an intermediate stage between the identification of the issues (previous research by Johnson et al, 2009) and the design of a capability database by means of a design-relevant national survey of capabilities (future research). The aim of this study is to provide a full feasibility study for the national survey. Since it is not feasible to use a range of actual product interactions in a large national survey, there is a need to identify the best proxy measures, i.e. the set of measures that best predicts product interaction. The findings of this stage will directly inform the choice of measures for designing a national survey for assessing the capabilities of a population sample. In essence, the findings from this stage will ensure that instructions, materials, tests, interview questions and response categories are set at the most appropriate levels for the proposed national capabilities survey.

2. Experimental methodology

Ethical approval was granted for the study from the Cambridge University Psychology Research Ethics Committee (Application No: 2009.26).

In preparation for the 100-person study, a smaller pilot was undertaken of 6 respondents to identify possible problems in running the full study. The objectives of the pre-pilot were to explore respondents' comprehension of specific modules, to time individual modules and to test some procedures such as the test protocols.

2.1 Sampling

A sample of 100 respondents aged 50-80 years was recruited from the Cambridge area, including both urban and rural respondents. Given the relatively small sample size of this intensive investigation, we decided to focus on the over 50 age group, where the range of variation in capabilities and product interactions would be expected to be large compared to a younger sample of the population.

The respondents were recruited through organisations such as the Cambridgeshire Older People's Enterprise and Age Concern; University of the Third Age in Cambridge; NHS Retirement Fellowship; and Over Sixties Clubs. In addition, flyers were posted on bulletin boards in libraries, supermarkets and other outlets. An incentive in the form of a £10 gift voucher was offered for participating in the study. Volunteers were asked to get in touch either through email or phone contact for an initial screening, which was conducted by phone.

Information about socio-demographics (such as age, gender, education and occupation) and health (with respect to vision, mobility, hearing and hand movements) was obtained at the initial screening. The aim was to ensure that the respondents were the desired age, represented the range of socio-demographic and health categories that would be expected in this age group and had no major mobility problems as access to the experiment room was limited. If the person met the inclusion criteria, at the end of the screening interview an appointment was allocated to attend the actual study at Addenbrooke's hospital. Of 109 volunteers, 8.3% were excluded after the screening interview.

The study involved one two-hour assessment session, which took place in a dedicated testing room. The respondents were asked to sign a consent form prior to taking part in the study.

2.2 Study design

Both self-reported and performance measures were used for the study. Self-reported socio-demographic information on age, gender, education, employment status, occupation, household income, living arrangements and ethnicity and information about technology use was collected. The study considered four domains of capabilities: vision, hearing, dexterity and cognitive function, since the vast majority of products make demands on these user capabilities.

The study included the following four modules: (1) component functions, (2) component activities, (3) psychological characteristics and (4) product interactions. Table 1 presents a summary of each module, populated by ticks to indicate whether self-report or performance test was conducted. In terms of the use of assistive devices we sought to replicate, as closely as possible, respondents' experience of everyday life. Specifically, the self-report questions required respondents to think about their past experience with no distinction made between tasks for which they use or do not use assistive devices. Similarly, for performance measures, respondents were allowed to use whatever assistive devices they had with them that they would normally use (e.g., eye glasses, hearing aid etc.).

[Table 1 near here]

2.2.1 Module 1: Component functions

Both self-reported measures and performance tests were used in this module for the assessments of vision, hearing and dexterity capabilities. Whereas, only performance tests were used for assessment of cognitive function capabilities.

<u>Vision</u>

Self-report measures – Respondents were asked questions relating to their reading and recognition capabilities: 'How good is your eyesight in general?', 'How good is your eyesight for seeing things at a distance, like recognising a friend across the street?' and 'How good is your eyesight for seeing things up close, like reading an ordinary newspaper print?'. The response options were 'excellent', 'very good', 'good', 'fair' and 'poor'. These questions were adapted from ELSA (English Longitudinal Study of Ageing) (Marmot et al, 2003). The ELSA study is based on the US Health and Retirement Study, which used some of the same measures.

Performance tests – Visual acuity and contrast sensitivity were measured using Test Chart 2000 (Thomson, 2005), a computerised comprehensive test chart system that presents optotypes on a computer monitor. Test Chart 2000 software was installed on a PC (Precision 340, Dell, Texas, USA) running a supplementary liquid crystal display (LCD) monitor (Acer, Taiwan) through a standard VGA connection.

a) Visual Acuity – Test Chart 2000 offers a number of visual acuity tests and the logMar chart was chosen for measuring visual acuity in this study because it is regarded as the gold standard of vision tests (Hazel and Elliot, 2002). In a logMar chart each row contains five letters and the letter size changes in steps of 0.1 logMAR between one row and the next. Based on the size of the experiment room, the logMar chart in the study was calibrated for 3.5m viewing distance. Therefore the scaled chart contained eleven rows with the letter sizes varying from 0.8 logMar (Snellen equivalent = 6/38) to -0.20 logMar (Snellen equivalent = 6/3.8). From 3.5m distance, the respondents read the chart, with the eleven rows shown simultaneously on the screen, starting from the biggest row. The endpoint for the visual acuity test was reached when at least three of the five letters in one row were read incorrectly. The visual acuity result was noted in Snellen (metres) format as it is easy to comprehend.

b) Contrast sensitivity – The Pelli-Robson chart (Pelli et al, 1988) was employed as the method of choice for measuring contrast sensitivity because it is the most frequently used chart in UK optometric practice (Latham, 1998; Thayaparan et al, 2007). The Test Chart 2000 contrast sensitivity test is similar to the Pelli–Robson test, where the chart is viewed at 1m distance and each triplet of letters also has the same contrast as the Pelli–Robson chart, although only one triplet is presented per line (Thayaparan et al, 2007). For the test, the chart was calibrated for the standard 1m viewing distance. Although the measurement conditions used were as standardised as possible, it should be noted that we were interested in relative differences in contrast sensitivity amongst our sample as a measure of predictive power, therefore the absolute value of measured contrast sensitivity for our respondents was not our criteria. The chart was left on for several minutes before testing to ensure that luminance had reached peak level. The mean screen luminance was 304cd/m^2, which was measured using a light meter. For the test, letters were displayed in triplets of decreasing contrast from the top to the bottom of the screen. From 1m distance, the respondents read the letters from top to bottom (with the interviewer scrolling down the screen to present each row) until they could no longer read two out of the three letters displayed. The respondent was assigned a score based on the contrast of the last group in which two or three letters were correctly read. The score was a measure of the respondent's log contrast sensitivity. The 'letter by letter' scoring system was used whereby each letter correctly identified was scored as 0.05 log units (except for the first triplet, where the contrast is 100%).

Hearing

Self-report measures – Respondents were asked questions related to their hearing status and speech discrimination capabilities: 'How good is your hearing in general?', 'How good is your ability to follow a conversation if there is background noise, such as TV, radio or children playing?' and 'How good is your ability to follow the dialogue in a movie or at the theatre?'. The response options were 'excellent', 'very good', 'good', 'fair' and 'poor'. The questions were adapted from ELSA (Marmot et al, 2003) and ONS Great Britain Disability Follow-up Survey 1996/97 (Grundy et al, 1999).

Performance tests - Hearing threshold was measured to determine functional hearing status. This was assessed with pure-tone air-conduction audiometry (Brender et al, 2006). Test equipment consisted of an Oscialla USB 300 Screening Audiometer, calibrated headphones with audiocup circumaural cushions to reduced ambient noise, and a response button. This hearing test uses a PC to present the softest (lowest volume) sounds that could be heard by the respondent at various frequencies. Respondents were correctly fitted with the headphones. Then they were asked to press a response button as soon as they heard a sound, even if very faint. For both ears separately, hearing thresholds were determined for tones of 0.5, 1, 2 and 4 kHz. Functional hearing status was determined on the basis of the pure tone average across 0.5, 1, 2 and 4 kHz (Dalton et al, 2003; Beria et al, 2007). This is considered a good estimate of an individual's hearing through the range of frequencies of speech (Voeks et al, 1993). The mean loss in decibels or pure tone average over the four frequencies was computed for each ear. The score of the 'better' ear was used.

Dexterity

Self-report measures – Respondents were asked five questions relating to functioning of their hands to perform certain tasks: 'How good is your ability to make fine finger movements?', 'How good is your ability to pick up small objects like a pin?', 'How good is your ability to tie things such as shoe laces or ribbon?', How good is your ability to grip and turn objects for example, gripping and turning a door handle in order to open a door?' and 'How often do you have difficulty in twisting the screw cap on a bottle in order to open it?'. The response options for the first four questions were 'excellent', 'very good', 'good', 'fair' and "poor". The response options for the fifth question were 'never', 'seldom', 'occasionally', 'frequently' and 'always'. Questions on picking up and tying objects were adapted from the ONS Great Britain Disability Follow-up Survey 1996/97 (Grundy et al, 1999) and the remaining questions were designed specifically for the purpose of the study to reflect everyday activities using hands.

Performance tests - Fine finger dexterity was measured using the Purdue Pegboard (Tiffin, 1948). This apparatus is a wooden pegboard with four cups for pins, collars and washers at the top of the board, and two columns of 25 holes each at the centre of the board. The assessment involves a series of 4 subtests of which the first three involve placing as many pins as possible into the holes with the right hand, then left hand, and then both hands – each in a 30-second period. In the last subtest the respondent uses alternate hands in order to make a specified series of assemblies consisting of pins, collars and washers, in a 60-second period. The score on the first two subtests is the number of pins inserted in the holes. The number of pairs of pins constitutes the score on the third subtest. The score on the assembly test consists of the number of assembled pins, collars and washers in the holes. (Note that out of the four subtests, only the assembly task was analysed for the purpose of the study, as this task involves the ability to integrate speed and precision with finely controlled discrete movements of the fingers and provides a good measure of fine finger dexterity.)

Cognitive function

Self-report was not employed for assessing cognitive function, since it is known to be an unreliable guide to cognitive performance, particularly memory (Huppert et al, 2006). For the performance measures, we selected domains of cognitive function which are of direct relevance to successful product interaction. These were: learning and memory; executive function (attention and speed of processing) and the basic skill of literacy. Unlike the domains of sensory and motor functions, there are no universal standards for how to assess cognitive function. Therefore we chose to use tests which have been used in other population studies of this age group. All the cognitive functions, except for the symbol learning test, were assessed by standard tests from ELSA (Steel et al, 2003a).

Learning and Memory:

a) Immediate and delayed memory – This concerns verbal learning and recall. A list of 10 words was presented aurally and the respondent was asked to recall as many words as possible immediately, and again after a short delay during which they performed another task (prospective memory task). The total number of words correctly recalled out of 10 provides a measure of immediate and delayed memory.
b) Prospective memory – This concerns memory for future actions and is sometimes referred to as ‘remembering to remember’. It was assessed by asking respondents to remember to carry out an instruction later in the session, namely, writing their initials at the top left hand corner of a work sheet when a clip board is handed over. A correct response requires the person to carry out the correct action without being reminded.
c) Symbol learning – This is not a standard test, but was developed for the purposes of the study. Some symbols and meanings were standard while others were novel. A list of 10 symbols and their corresponding meaning was presented to the respondents for 30 seconds and they were asked to memorise them. The respondent was then shown the symbols again and asked to recall the meanings of as many symbols as possible. The total number of meanings recalled correctly out of 10 provides a measure of symbol learning.

Executive function:

d) Cognitive speed and accuracy - Assessed using a visual search task involving letter cancellation. The respondents were asked to cross out as many target letters (P & W) as possible in 1 minute on a page of 780 letters arranged in rows of 30 letters. The total number of letters searched gives a measure of speed of processing. The number of target letters missed up to the letter reached by the respondent provides a measure of accuracy.

Basic skill:

e) Literacy – A large print medicine label describing instructions for taking a medicine was given to respondents to read and was followed by asking four questions to establish how well the respondents understood the instructions on the label. The literacy test score is the total number of correct answers and the maximum score on this test is 4.

2.2.2 Module 2: Component activities

Ability to engage in key activities involved in product interactions is measured in this module using both self-report questions and measures of actual performance.

The study design involved activities which are commonly used when interacting with a range of products. The following component activities were chosen for the study:

A. Reading

i) *Reading text on a digital display* – This task involves reading text written on the display of a mobile phone at three different luminance levels – high, medium and low.

ii) *Reading text on a plastic surface* – This task involves reading black text (2.5mm height) on the cream background of a shampoo bottle.

B. Speech recognition

i) *Hearing messages at different volumes* – This task involves listening to pre-recorded statements in a mobile phone at different volumes (high, medium and low) and repeating them out-loud.

C. Manual co-ordination

i) *Twisting a dial on a kitchen timer* – This task involves setting a kitchen timer to 45 minutes by twisting and turning.

ii) *Using a touch screen keypad* – This task involves typing a ten-digit number using the touch screen keypad of the mobile phone.

iii) *Pressing buttons on a remote control* – This task involves locating and pressing the PLAY button on the remote control of a DVD player.

Self-report measures - Self-reported information on the respondent's frequency of product use and difficulty in product use was collected for each type of component activity. If the type of product had not been used before, we asked about expected difficulty with the product interaction. The response options for frequency questions were 'never', 'less than once a month', 'once a month', 'more than once a month', 'once a week', 'at least three times a week', 'once a day' and 'more than once a day'. The response options for difficulty questions varied from 'extremely easy' through 'neither hard nor easy' to 'extremely difficult'.

Performance tests - Performance involved scoring whether or not the activity was successfully completed.

In addition to the specific component activities related to product interaction, information on more global functioning was obtained by self-report with questions regarding general instrumental activities of daily living (IADL). The IADL addressed the following activities (Lawton and Broody, 1969): using a map to figure out how to get around in a strange place; preparing a hot meal; shopping for groceries; making telephone calls; taking medications; doing work around the house or garden, and managing money, such as paying bills and keeping track of expenses. Respondents were asked whether or not they had any difficulty doing each of these activities within the past month.

2.2.3 Module 3: Psychological characteristics

<u>General</u>

Respondents were asked questions about their self-efficacy, self-esteem, optimism and mastery. An example of each of the four measures include: 'I can solve most problems if I invest the necessary effort'; 'In general I would describe myself as a confident person'; 'Overall I expect more good things to happen to me than bad' and 'I feel like giving up quickly when things go wrong'. The questions for self-efficacy and mastery used in our study were adapted from the General Self-Efficacy Scale (Schwarzer, 1993), self-esteem questions were adapted from the Rosenberg Self-Esteem Scale (Crandal, 1973) and optimism items from the Life Orientation Test (Scheier et al, 1994). Each of the four measures was rated on a 7-point scale ranging from 1 (disagree strongly) to 7 (agree strongly).

Product-related

Standard measures were not available so we designed the questions for product-related self-efficacy, self-esteem, optimism and mastery. An example of each of the four measures include: ‘I am confident in my current skills and ability to learn how to use a new piece of equipment in my job’; ‘If I experience difficulty using a piece of equipment I feel incompetent and blame myself’; ‘With a new piece of equipment I usually expect that I will be able to use it without too much difficulty’ and ‘When I get a new piece of equipment, if it’s not obvious how to use it, I give up’. Each of the four measures was rated on a 7-point scale ranging from 1 (disagree strongly) to 7 (agree strongly).

Personality

‘The Big Five’ personality characteristics (extraversion; openness to new experiences; emotional stability; conscientiousness and agreeableness) were measured using the Ten-Item Personality Inventory (Gosling et al, 2003), which is the most widely used model. Each of the five personality characteristics was rated on a 7-point scale ranging from 1 (disagree strongly) to 7 (agree strongly).

2.2.4 Module 4: Product interactions

Ability to interact with a range of household products was measured in this module using both self-report questions and measures of actual performance. Time was invested in market research which explored various types of common household products. The models of the products chosen for the study had typical features and functions, and were assessed as being of average difficulty to use, relative to the competing models. A typical iron (Philips GC2522), microwave (Panasonic NN-E255WB) and a landline telephone (BT Décor 1200) were selected to assess human-product interaction on the basis of five primary considerations: (a) products that are familiar and relatively easy to use, (b) the products between them made demands on the four main capability domains (vision, hearing, dexterity and cognitive function), (c) use of each product heavily relies on at least two of the four domains, (d) selection of the performance tests for each product interaction considered floor (too many people failing) and ceiling effects (too many obtaining maximum scores) and (e) the selected performance tests did not require reading the manual.

Initially, a digital radio (Pure Tempus-1S) was also chosen, but was removed after the initial pre-piloting stage because no-one succeeded in carrying out apparently simple tasks such as presetting a favourite radio station and setting an alarm.

Self-report measures - Self-reported information on the respondent’s frequency of product use and difficulty in product use was collected for each type of product, using the same structure as the component activities module.

Performance tests - Two tasks were given for each product to assess product interaction. Measures involved scoring of the time taken to complete the specified task and whether or not the task was successfully completed.

A. Iron

Task 1: *Fill water tank* – This task involves opening the cap of the filling opening, tilting the iron and filling the water tank up to maximum level and finally closing the cap of the filling opening.

Task 2: *Set dial to cotton* – This task involves putting the iron on its heel and turning the temperature dial to cotton/linen by making sure that the dial was aligned with the white mark.

B. Microwave

Task 1: *Heat food at high temperature for 2 minutes 30 seconds* – This task involves pressing (i) the micro-power button to select the desired power level, (ii) the minutes and seconds buttons to set the desired time (i.e. pressing the minutes button twice and the seconds button three times to set the time to 2 minutes and 30 seconds) and finally (iii) the start button.

Task 2: *Defrost chicken weighing 230 grams* – This task involves pressing the auto-defrost button twice to get to 'meat items', setting weight using the up and down buttons to 230 grams and then finally pressing start button (time is pre-set).

C. Landline telephone

Task 1: *Set ringer volume* – This involves pressing (i) the menu button to display ringer volume option, (ii) the tick button to select the ringer volume option, (iii) the redial or calls button to change the volume level, (iv) the tick button to confirm and save the volume desired and (v) the cross button to return to standby.

Task 2: *Store phone number* – This task involves pressing the directory button, pressing the tick button to display Add New Entry option, entering the name for the entry using keypad and then pressing the tick button, entering the number using keypad and then pressing the tick button and finally the cross button to return to standby.

Socio-demographics

Background information was collected on the socio-demographic characteristics of the sample, using self-report measures. Respondents were asked about their age, gender, the highest educational qualifications they attained, employment status, household income, whether anyone lives with them in the household and ethnicity. For these questions, respondents were given a pre-determined list of response options to choose from. Respondents were also asked about their main occupation. Based on this information, a three-category occupational classification (managerial and professional; intermediate; manual and routine) was derived. The questions and response options for the socio-demographic characteristics were adapted from ELSA (Marmot et al, 2003).

Technology use

Respondents were asked which of the following technology items or services they own or have access to for personal use: personal computer, internet, telephone, mobile phone, email, wii, CD player, DVD player, digital/satellite TV.

Additional features of the study

Care was taken to control for potential order effects by counter-balancing across respondents the order in which the four modules were performed and randomly assigning the order of tasks within the product interactions and component activities modules.

2.3 Statistical analysis

The statistical program SPSS (Statistical Package for the Social Sciences) 16.0 for Windows was used to analyse the data. Spearman's rank correlation coefficients were calculated to examine the relationship between self-reported and performance measures. Spearman's rank correlation coefficients were used instead of Pearson's correlation coefficients because non-linear, as well as linear, associations between the variables were of interest; and the Spearman's rank correlation coefficient does not rely on any distributional assumptions. The initial descriptive statistics for the three distinct types of predictor variables, (a) component functions, (b) component activities and (c) psychological characteristics, were stratified by age and examined, in order to confirm the expected age-related variation. For examining the age-related variation, the sample was categorised into a younger group (50-64 years) and an older group (65-80 years). Having measures which fit the expected pattern of a decline with age is a necessary (though not sufficient) condition for being a useful predictor of capability. Multiple logistic regression analysis was undertaken to assess which component functions, component activities and psychological characteristics provided the most accurate and generalisable prediction of success or difficulty in the product interactions. The comparative predictive power of the following types of potential predictive variables was analysed using a multiple logistic regression analysis approach: (i) self-reported component functions, (ii) measured ability on component functions, (iii) self-reported component activities, (iv) measured ability on component activities and (v) self-reported psychological characteristics.

3. Results

3.1 Socio-demographic characteristics of the sample

The basic description of the characteristics of the sample is presented in Table 2. In the study, 56% of respondents were women as would be expected in a population above the age of 50, where women outlive men (Gjonca and Calderwood, 2003). The mean age of the sample was 67 (SD = 5.9). Two-thirds (66%) were aged 65 and above. 51% had achieved a high educational qualification (degree or equivalent) and 49% achieved an intermediate education or no qualifications. The largest proportion of the sample (72%) lives with others (children or spouse/partner or both) and only 28% of the sample live alone. In total, 88% of the sample was currently unemployed (i.e. retired or unable to work) and only 12% were currently in employment (part-time or full-time). 50% of the sample were in professional or managerial positions, 31% in intermediate positions and 19% were in manual or routine occupations. The largest proportion of the sample (77%) had a household income in the range £10,000-£50,000 with a higher proportion of younger respondents in this income category (91%) compared to the older respondents (68%). With regard to the ethnic composition, the 'white' category was by far the largest (98%).

[Table 2 near here]

In response to the technology use questions, 82% of the sample used 6 or more of these items or services, with little variation between the younger (88%) and older (79%) age groups, indicating that most of the sample was inclined towards using technology.

3.2 Summary statistics

Tables 3 to 6 present the summary statistics based on the self-reported and performance measures for the sample including means, standard deviations and percentages by age group for the four modules. To avoid the problem of spurious results related to multiple testing, we didn’t conduct individual tests of statistical significance, but rather included age as a factor in the regression analysis (see section 3.4).

3.2.1 Module 1: Component functions

As expected, on all the vision questions, a higher proportion of people in the older group rated their vision as fair or poor (rather than excellent, very good or good), compared to the younger group. Measured visual acuity also decreased with age from a mean value of 6/6.6 in the younger group to 6/7.5 in the older group. A visual acuity score of 6/6 is considered normal vision and acuity scores beyond 6/6 signify poorer vision (Elliot et al, 1995; Hirvela and Laatikainen, 1995; Tate et al, 2005). It is interesting to note that the older age group performed better than the younger ones in the contrast sensitivity test (2.4 vs 1.9 log units). In general, a Pelli-Robson score of 2.0 indicates normal contrast sensitivity and scores less than 2.0 signify poorer contrast sensitivity (Hohberger et al, 2007). The mean contrast sensitivity value for the younger group was only marginally less than 2.0, implying that their ability to see low-contrast objects under conditions of poor visibility was close to normal.

[Table 3 near here]

A higher proportion of people in the older group reported their hearing as fair or poor on all the hearing questions. Measured hearing also decreased with age, where the older age group (hearing range varied from -10 to 52.5 dB) had higher mean pure tone average than the younger group (hearing range varied from 1.25 to 28.8 dB). For the hearing test, higher score implies poorer performance. More than half (67%) of the sample had normal hearing and 33% were impaired. Of those impaired, 88% were over 65 years old.

On most of the dexterity questions, as expected, a higher proportion of people in the older group rated their dexterity as fair or poor. However for questions relating to tying objects and turning objects, the results were not as expected, with younger respondents reporting less ability on these tasks. These findings should be further explored in future work. With regard to the dexterity performance test, fine finger dexterity was considered normal if the total number of assembled pins, collars and washers in the holes is equal to 28 or more. The younger group placed more pins, collars and washers (average score of 29.6) than the older group (average score of 26.8) in this test.
The older group performed slightly worse than the younger group on all the cognitive tests. They remembered fewer symbol meanings (6.4 vs. 8.3) and fewer words immediately (6.1 vs. 6.9) and after a delay (5.2 vs. 5.9). They were outperformed by the younger group (60.6 % vs. 70.6%) in successfully carrying out the task without a reminder in the prospective memory test. They performed slightly slower and made slightly more errors in the letter cancellation task. However, there was no age difference on the literacy test.

3.2.2 Module 2: Component activities

Table 4 shows the mean self-report ratings of frequency of use and difficulty for each component activity, and successful completion of each task. On all the products (expect for the kitchen timer and the remote control), a higher proportion of people in the younger group rated their frequency of use as high and difficulty of use as low, compared to the older group. The mean ratings for frequency and difficulty of use of a kitchen timer are the same for both the age groups. For the remote control, a higher proportion of people in the older group rated their frequency of use as high and their difficulty of use as high, compared to the younger group. As expected, the older group performed worse on almost all of the reading, speech recognition and manual co-ordination activities.

Respondents who reported not being able to perform one or more instrumental activities of daily living (IADLs) were considered to have impaired IADL. The prevalence of reported difficulty with IADLs increased with age from 44.1% in respondents aged 50-64 to 57.6% of those aged 65-80 (data not presented).

[Table 4 near here]

3.2.3 Module 3: Psychological characteristics

Table 5 compares respondents' general approach to life with their approach to product use. Higher scores indicate higher levels of the measured concept. The average score on general self-efficacy was the same (5.4) for both age groups. The average scores on general self-esteem and general optimism decreased with age, from 5.0 and 5.3 respectively in the younger group to 4.8 and 4.9 in the older group. It is interesting to note that the older group has a higher mean score than the younger group on general mastery – an average of 5.3 versus 4.8. There is a slight difference in the average scores between younger and older groups on product self-efficacy and product optimism. The average score on product self-esteem was the same (4.5) for both age groups. The younger group had a higher average score (5.6) on product mastery than the older group (5.4). On comparison of personality characteristics, the older group perhaps surprisingly had higher average scores on openness to new experiences, extraversion and emotional stability. On the other hand, the younger group had a higher average score on conscientiousness and agreeableness.

[Table 5 near here]

Spearman correlations were calculated to examine the relationship between general and product-related items. Correlations were higher for self-efficacy and mastery than for optimism and self-esteem. The correlations were 0.44 ($p<0.001$) between general self-efficacy and product self-efficacy; 0.46 ($p<0.001$) between general mastery and product mastery; 0.31 ($p=0.002$) between general optimism and product optimism; and 0.21 ($p=0.036$) between general self-esteem and product self-esteem.

3.2.4 Module 4: Product interactions

As can be seen in Table 6, the frequency of use of a landline telephone and on iron decreased with age. In contrast, frequency of use of the microwave increased with age. On all the three products, a slightly higher proportion of people in the older group rated higher difficulty in product use. As expected, on all the product interaction tasks a higher proportion of people in the younger group completed the tasks successfully. Of all the tasks, the poorest performance was seen on the second microwave task (only 9% of respondents were successful) and the first landline telephone task (only 15% of respondents were successful). The poor performance on the second microwave task may be a result of it being a fairly unusual task. It is not known how often people actually use the 'auto-defrost' function on microwaves rather than approximating the task by, for example, heating the food for a short time at a particular power level. The poor performance on the first landline telephone task may be the result of respondents not knowing that they need to access the 'menu' function to find the ringer volume option. In terms of the time taken to successfully complete tasks, variations appeared across age groups and tasks. For example, the older group took longer than the younger group to successfully complete each of the iron tasks, while the younger group took longer than the older group on the first microwave task and the first landline task.

[Table 6 near here]

3.3 Correlation between self-report and performance measures

Table 7 shows the correlations between self-reported and performance measures. The strength of this association for physical functioning (vision, hearing and dexterity) is at best only moderate. Correlations between self-report and performance measures were higher for dexterity and hearing than for vision. For component activities, very weak associations between self-report and performance measures were found for most of the activities, except for two whose values were moderate. Self-reported difficulty in using a touch screen was moderately correlated with actual performance. Similarly self-reported difficulty in twisting dials was moderately correlated with the actual performance of twisting a kitchen timer. With regard to product interactions, very weak associations were found between self-report and performance measures.

[Table 7 near here]

3.4 Significant predictors of successful product interaction

A multiple logistic regression analysis approach was adopted to investigate which measures best predict successful product interaction. It should be noted that the analysis approach used for the study was exploratory. That is, the analysis should not be interpreted as a definitive investigation but rather an evaluation of which component functions, component activities, and psychological characteristics will contribute to the prediction of successful product interactions.

For successful product interaction, the ‘task success’ variable (whether or not the task was successfully completed) and not the ‘time taken’ variable was used as an outcome measure for each product task. The fundamental problem was that on its own, ‘time taken’ is not very useful because a short time could be (a) due to the respondent completing the task quickly or (b) due to the respondent finding they cannot do it at all and giving up quickly. Therefore, the only easy way to include this variable in the analysis was to look at time taken within the successful completion group. However, the reduced sample size that would have resulted from this step would have resulted in insufficient statistical power.

Three outcome measures were used in the analysis: (i) for the iron, we chose ‘both iron tasks completed’ as an outcome measure because a large number of respondents were able to complete both the tasks, (ii) for microwave and (iii) landline telephone we chose ‘at least one task completed’ as the outcome measure because very few respondents were able to complete both tasks. Three distinct types of predictor variables were used: (a) *component functions*, (b) *component activities* and (c) *psychological characteristics*.

As a first step, multiple regression models were used to find significant predictors of each outcome measure among the socio-demographic and technology use variables. Forwards model selection with likelihood ratio testing was the method of choice, whereby these variables were added to the model in turn and only remained if they gave a significant improvement to the goodness-of-fit of the model. A 5% significance level was used throughout.

Secondly, multiple logistic regression models were used with forwards model selection (likelihood ratio testing) to find significant predictors of each outcome measure among the *component functions* variables. This consisted of applying separate logistic regressions to each of four sets of predictor variables: (i) self-report measures of vision, hearing and dexterity, (ii) performance measures of vision, hearing and dexterity, (iii) both self-report and performance measures of vision, hearing and dexterity combined, and (iv) performance measures of cognitive function.

Thirdly, multiple logistic regression models were again used to find significant predictors of each outcome measure among the *component activities* variables. This consisted of applying separate logistic regressions to each of four sets of predictor variables: (i) self-report measures of 6 *component activities*, (ii) performance measures of 6 *component activities*, (iii) both self-report and performance variables combined, and (iv) the general IADL variables.

Fourthly, multiple logistic regression models were used to find significant predictors among the *psychological characteristics* variables. This consisted of applying separate logistic regressions to each of four sets of predictor variables: (i) general psychological characteristics, (ii) product-related psychological characteristics, (iii) both general and product-related psychological characteristics combined, and finally, (iv) general psychological characteristics, product-related psychological characteristics and personality variables combined.

Table 8 shows the results of the first four steps, along with the odds ratios (OR) and 95% confidence intervals (CI). Odds ratios provide a way of comparing groups according to the odds of outcome within each group. (In this case, the odds of outcome is the proportion of individuals with successful task completion over the proportion without successful task completion.) A 95% confidence interval of an odds ratio shows the interval within which the true ratio of odds in the population is likely to lie with 95% probability. The table presents those variables which were significant in the final model after forwards model selection at the 5% level of significance (likelihood ratio testing). From this, we identified the individual variables which were significantly associated with the product interaction outcomes. For instance, in *component functions*, out of the vision, hearing and dexterity performance measures, visual contrast sensitivity was a significant predictor of successfully completing one of the microwave tasks.

[Table 8 near here]

The initial results of separate logistic regression analyses on each category of predictor variables show that within each category there were significant predictors of product interactions. Among socio-demographic variables, age, gender, household income and occupation were significant predictors of product interactions. In the case of *component functions*, two self-report measures of vision (general vision, close vision), no self-report measures of hearing and two self-report measures of dexterity (finger movements, twisting) were significant predictors of product interactions, as were two performance measures of vision (visual acuity and contrast sensitivity), one performance measure of hearing (hearing threshold), no performance measures of dexterity and three performance measures of cognitive function (memory recall, prospective memory and speed of processing). With respect to *component activities*, three self-report measures (frequency of reading text on digital screen, difficulty of reading text on digital screen, difficulty in twisting dials) and three performance measures (successful task completion of pressing buttons, successful task completion of hearing at low volume, successful task completion of using touch screen) were significant predictors of product interactions. The *psychological characteristics* which proved to be significant predictors were general self-efficacy, general mastery, product self-efficacy and openness to new experiences.

For the final step, all variables which were found to be significant in the previous analyses were then entered into a set of analyses – one regression analysis for each product interaction. Forwards model selection was used to identify the strongest predictors of product interaction. A 1% significance level with 99% confidence intervals was used for this final set of analyses in order to restrict the numbers of variables in the final models to the most important. Table 9 shows the results, along with the odds ratios (OR) and 99% confidence intervals (CI). For instance for the microwave, contrast sensitivity was found to be a significant predictor (OR 409.83, 99% CI 1.63 to 103016). Higher values of contrast sensitivity correspond to a higher probability of successful task completion. However, this regression model exhibited signs of instability due to the small sample size; hence, the extremely wide confidence intervals.

[Table 9 near here]

The final results of the multiple logistic regression analysis (which took the significant variables from the initial multiple logistic regression analysis) for identifying the strongest predictors show that in the case of *component functions*, self-report measures of vision, hearing and dexterity were not good predictors of product interactions. Whereas, two performance measures of vision (visual acuity, contrast sensitivity), one performance measure of hearing (hearing threshold) and none of the performance measures of dexterity and cognitive functions were significant predictors. With respect to *component activities*, one self-report measure (frequency of reading text on digital screen) and one performance measure (successful task completion of hearing at low volume) were significant predictors of product interactions. Only one self-reported *psychological characteristic* (openness to new experiences) was a significant predictor.

4. Discussion

This multi-dimensional experimental study was designed as a precursor to a major national survey of capabilities, with the longer-term aim of producing data which will help designers make more inclusive design decisions. Since a large-scale national survey conducted in people's homes could not realistically measure performance on a range of human-product interactions, we have designed the present study as an intermediate step. Its objective was to establish the extent to which component functions, component activities and psychological characteristics could predict human-product interactions and hence be used as proxies in a future national survey. Previous influential surveys such as the ONS Great Britain Disability Follow-up Survey relied entirely on self-report assessment of capabilities, and a subsidiary aim of the present study was to compare the effectiveness of self-report versus performance measures as predictors of product interactions.

For this study, we selected four domains of capabilities which are involved in the great majority of product interactions, namely: vision, hearing, dexterity, and cognitive function.

We selected three common household products for which successful performance relies on combinations of these capabilities. The products were an iron, a microwave, and a landline telephone, selected to be typical products of their kind and average in terms of price and ease of use.

Main findings

a) Age differences on self-report and performance measures

It is common to find a decline in capability with increasing age. An important initial step in assessing the validity of the measures used in the current study is confirming the presence of this pattern. We found that the prevalence of vision, hearing, dexterity and cognitive problems increased with age in our sample. Our findings are in line with other studies, which show significant age-related differences in self-reported vision and performance on vision tests (Tate et al, 2005; Charles, 2007). It is worth noting that our study found that 7% of our sample had fair or poor vision, which corresponds exactly to the results of a previous national survey, ELSA Wave 2 (Steel et al, 2003b). Likewise, our findings on hearing from both self-report and performance measures, are in line with existing studies (Weinstein and Ventry, 1982; and Voeks et al, 1993). Our study found that 43.9% of respondents aged over 65 years reporting difficulty following a conversation if there was background noise, which is very similar to 45% reported by ELSA Wave 2 (Steel et al, 2003b). The percentage of people who had hearing loss in our sample was 33%, which falls in the low end of range of hearing loss (31% to 87%) reported from other studies (Harford and Dodds, 1982; Weinstein and Ventry, 1982; Thomas et al, 1983 and Voeks et al, 1993). Regarding dexterity, age and fine dexterity measured by the Purdue Pegboard are closely and inversely related, which accords with previous studies (Desrosiers et al, 1995; Pennathur et al, 2003). Our findings are in line with other studies which show age-related decline in memory performance on the tests of recall (Huppert et al, 2006) and prospective memory (Huppert et al, 2001) and speed at which one can process information (Salthouse, 1991). If the findings on prospective memory (that is, remembering to carry out a task without being reminded) are indicative of forgetfulness in daily life, then the high prevalence of age-associated forgetfulness in the sample is a cause for concern, particularly in the oldest age group (Huppert et al, 2001). These findings raise questions about the extent to which older individuals remember to carry out essential actions such as those concerned with health (taking medication), security (locking doors, turning off the cooker) and economic activity (collecting pensions, checking statements).

Age differences were also observed on all six performance measures of component activities. To the best of our knowledge, this is the first evidence for age-related differences in activities that comprise common components or sub-components of product interactions. This finding is important because it represents an intermediate stage between basic sensory and motor functioning and success with interacting on common household products. There was no consistent pattern of age difference with respect to psychological characteristics. Our results from regression analysis showed that age predicted successful performance on all three product interactions.

b) Self-report versus performance measures

The results show that the strength of association on most of the domains examined in this study is moderate or poor. Within the domains of dexterity and hearing, stronger associations were found on average compared to vision. These results support other studies (Gatehouse, 1991; Myers et al, 1993; Kempen et al, 1996). One exception is a study by Bergman and

Sjostrand (1992), which showed relatively a stronger correlation between self-report and performance measures for vision, but the sample was very elderly (mean age 82 years). Further investigation is required to identify factors responsible for discrepancies between self-report and performance measures in sensory and motor functioning. In general, the correlations for component activities and product interactions were either low or non-significant, except for two component activities. The exceptions are moderate correlations between self-reported difficulty in using a touch screen and actual performance, and self-reported difficulty in twisting dials and the actual performance. These findings strongly suggest that a database of capabilities relevant to design cannot rely on self-report measures alone. Thus, data from ONS Great Britain Disability Follow-up Survey 1996/97 should be treated with caution, since it used only self-report measures. The relatively low correlations between self-report and performance measures suggest that they may be measuring different aspects of functioning, and thus produce different but valuable information.

c) Significant predictors of successful product interaction

As a general observation, for those significant variables identified by the initial regression analyses on each category of predictor variables (component functions, component activities and psychological resources), the odds ratios appeared to be generally higher for the component functions and component activities than for psychological characteristics. In our sample, psychological characteristics appear to be less important predictors of successful product interactions. Indeed, only one psychological characteristic (openness to new experiences) was found to be a significant predictor for only one of the outcome measures (landline telephone) in the final regression analysis. However, this observation needs to be confirmed in a larger sample.

The initial and final regression analyses showed quite different results for each of the product interactions (iron, microwave and landline telephone) in so far as the lists of significant predictors were fairly different. This suggests that the results appear to depend quite strongly on the type of product used for product interaction. Yet despite this, some factors did appear in more than one product domain in the initial regression analysis e.g. visual acuity, prospective memory, self-reported frequency of reading text on a digital screen, successful task completion of hearing at low volume, general self-efficacy, and product self-efficacy. Indeed, self-reported frequency of reading text on a digital screen was identified as a significant predictor in the final regression analysis for both the iron and landline telephone, indicating that this variable may have more general applicability as a useful predictor or proxy in future studies. It is not clear from our results if self-reported measures or performance measures are better predictors of product interaction. However, there is no reason to discount self-report measures as potentially useful predictors in a national survey.

Strengths, limitations and future work

This research represents the first attempt to integrate a wide range of design-relevant capability measures into a single large study. It had two principal aims:

1. To integrate self-report and performance measures across a variety of component functions, component activities and psychological characteristics as well as a series of actual product interactions.
2. To assess how well each type of measure can predict people's capability when interacting with products.

In addition to these methodological innovations, we used a sample which was reasonably typical of a mature population (50-80 years) and a study design that reduced the likelihood of undesirable order effects by randomising the sequence in which the modules were presented to the respondents.

The multiple logistic regression approach was adopted in the spirit of exploratory data analysis by conducting initially separate regression analysis on the three distinct types of predictor variables (component functions, component activities, psychological characteristics) and then combining significant variables into a final analysis for identifying the strongest predictors of product interaction. There are however, some obvious limitations to these results. The sample size of 100 was relatively low for such a large number of regression analyses, therefore there was low power to detect significant differences/associations. Accordingly, some potentially meaningful associations may have been missed (Type 2 error). A large number of regression analyses have been performed which means that the Type 1 error rate was likely to have been high (that is, some of the associations found may have been due to chance). However, to balance this issue, a more stringent level of significance was used in the final analysis (1%).

Our outcome measure of product interaction focussed on 'task success', but it would be good in future studies to also consider 'time taken' to perform a task successfully. While considering predictors of the actual performance on 'product interactions' as the dependent variable provides extremely interesting information, it is unlikely that a single factor will have strong predictive power over this variable, because the tasks that were set for these product interactions involved so many different aspects of user abilities. It would be interesting to consider, in the future, the actual performance on the 'component activities' as the dependent variable, and looking at the factors that contribute to predicting these, such as self-reported frequency of use or difficulty on the component activity, as well as self-report and performance measures of component functions related to the component activities.

Due to time limitations, the study considered only four domains of user capabilities: vision, hearing, dexterity and cognitive function. It will be important that future studies consider additional capabilities such as mobility, locomotion, reach and stretch. The study was conducted in an experimental room rather than in the respondent's home, which could have influenced our results, where respondents may have pushed themselves during a performance test in a way that they would not be able to do on a regular basis (Johnson et al, 2009). The sample includes older adults only, hence we may not be able to generalise our results to other age groups. Also, the sample was made up of volunteers, who are quite likely to be different in terms of capability from those who did not want to or could not volunteer. Similarly, the sample that responded to our specific recruitment methods may differ from the population we were not able to reach or to those who chose not to respond to our request for participants. Finally, limitations in terms of the space available for conducting the study required that we excluded those with mobility problems. While these factors mean that our sample was not truly representative of the population, we believe that the sample was suitable given the exploratory nature of this study. It will be important that future research is conducted with a broader, more fully representative sample.

Based on the results, items can be eliminated from the subsequent design-relevant national survey of capability on the basis that they are unlikely to be good predictors of product interaction. That having been said, it was recommended that some be retained on the basis of their face validity. For instance, Laitinen et al (2005) recommend including self-assessment of visual function along with the performance-based visual function tests in ophthalmologic studies, despite the modest correlation between self-reported and performance-based visual function tests (Hiller and Krueger, 1983; Carta et al, 1998; Rubin et al, 2001). This is because the respondent's own assessment of visual function may give additional information compared with visual acuity tests, which do not cover all visual components such as visual field, which influences the individual's ability to cope with his or her environment. Keeping this in view, we recommend that self-report ratings of capability in component functions (e.g. self-reported rating of vision, hearing, dexterity and other domains of human capabilities) be retained for the subsequent survey.

5. Conclusion

The findings of the research presented in this paper will inform the choice of measures for designing a national survey for assessing the capabilities of people. In essence, the findings provide valuable information on the appropriate levels of instructions, materials, interview questions, tests and response categories to be used in a future national capability survey. As an essential step towards this goal, a national pilot survey is being undertaken across a representative UK population sample of 400 adults by considering all the major domains of human capabilities (vision, hearing, dexterity, reach and stretch, locomotion, cognitive function), for building the capability database. Ultimately, the resulting database, in tandem with tools that present the data to designers in a relevant and accessible way, should promote and facilitate Inclusive Design.

Acknowledgements

This research was supported by an Engineering and Physical Sciences Research Council grant (Award Number: RG45089), which funded or part-funded three of the co-authors (RT, DJ and LG). We wish to express gratitude to the study participants and to our consortium partners from Engineering Design Centre (Cambridge), Royal College of Art (London), and Ergonomics and Safety Research Institute (Loughborough) for interesting discussions in preparation for this study.

References

Bajekal, M., Breman, R., Harries, T. & Woodfield, K., 2004. Review of disability estimates and definitions: A study carried out on behalf of the Department for Work and Pensions. Department for Work and Pensions, London.

Bergman, B., and Sjostrand, J., 1992. Vision and visual disability in the daily life of a representative population sample aged 82 years. Acta Ophthalmologica, 70 (1): 33-43.

Beria, J. U., Raymann, B. C., Gigante, L. P., Figueiredo, A. C., Jotz, G., Roithman, R., 2007. Hearing impairment and socioeconomic factors: a population-based survey of an urban locality in southern Brazil. Revista Panamericana de Salud Pública, 21(6): 381–387.

Brender, E., Alison, E.B., and Richard, M.G., 2006. Audiometry, The Journal of American Medical Association (JAMA), 295(4): 460.

Carta, A., Braccio, L., Belpoliti, M., Soliani, L., Sartore, F., Gandolfi, S.A., Maraini, G., 1998. Self-assessment of the quality of vision: association of questionnaire score with objective clinical test, Current Eye Research, 17:506 –11.

Charles, N., 2007. Estimates of the number of older people with a visual impairment in the UK, British Journal of Visual Impairment, 25(3), 199 - 215.

Clarkson, P.J., Coleman, R., Hosking, I., Waller, S., 2007. Inclusive Design Toolkit, Engineering Design Centre, University of Cambridge.

Coleman, R., 2001. Designing for our future selves. In: W.F.E. Preiser and E. Ostroff (eds) Universal design handbook. NewYork: MacGraw-Hill, 4.1–4.25.

Crandal, R., 1973. The measurement of self-esteem and related constructs. In: J.P. Robinson & P.R. Shaver (Eds), Measures of social psychological attitudes. Revised edition. Ann Arbor: ISR, 80-82.

Dalton, D. S., Cruickshanks, K. J., Klein, B. E., Klein, R., Wiley, T. L., & Nondahl, D. M., 2003. The impact of hearing loss on quality of life in older adults, Gerontologist, 43: 661-668.

Desrosiers, J., Hebert, R., Bravo, G., Dutil, E., 1995. The Purdue pegboard test: normative data for people aged 60 and over Disability and Rehabilitation, 17(5): 217-224.

Elliott, D.B., Yang, K.C., and Whitaker, D., 1995. Visual acuity changes throughout adulthood in normal, healthy eyes: seeing beyond 6/6, Optometry and Vision Science, 72(3): 186–191.

Elton, E**.,** and Nicolle, C., 2010. The Importance of context in Inclusive Design. Ergonomics Society Annual Conference 2010 - Keele University. Contemporary Ergonomics 2010.

European Commission (EC), 2008. European health interview and health examination surveys database [online]. European Commission. Available from: https://hishes.iph.fgov.be/index.php?hishes=home [Accessed 13 August 2008].

Forgas, J., 2008. Affect and Cognition. Perspectives on Psychological Science, 3**,** 94-101.

Gatehouse S., 1991. The role of non-auditory factors in measured and self-reported disability. Ada Otolaryngol, 476 suppl: 249-56.

Gjonca, E., and Calderwood, L., 2003. Socio-demographic characteristics. In: Marmot et al (Eds) Health, wealth and lifestyles of older population in England: The 2002 English Longitudinal Studies of Ageing. IFS. London.

Gosling, S.D., Rentfrow, P.J., and Swann, W.B., Jr. 2003. A Very Brief Measure of the Big Five Personality Domains, Journal of Research in Personality, 37: 504-528.

Grundy, E., Ahlburg, D., Ali, M., Breeze, E., Sloggett, A., 1999. Research report 94 – Disability in Great Britian. Corporate Document Services, London, ISBN 1-84123-119-3.

Harford, E.R., and Dodds, E., 1982. Hearing status of ambulatory senior citizens. Ear Hear, 3: 105-9.

Hazel, C., and Elliot, D., 2002. The dependency of logMAR visual acuity measurement on chart design and scoring rule, Optometry and vision science, 79(12): 788-792.

Hiller, R., and Krueger, D.E., 1983. Validity of a survey question as a measure of visual acuity impairment, American Journal of Public Health, 73: 93– 6.

Hirvela, H., and Laatikainen, L., 1995. Visual acuity in a population aged 70 years or older; prevalence and causes of visual impairment, Acta ophthalmologica Scandinavica, 73: 99 –104.

Hohberger, B., Laemmer, R., Adler, W., Juenemann, A., and Horn, F., 2007. Measuring contrast sensitivity in normal subjects with OPTEC 6500: influence of age and glare, Graefe's Archive for Clinical and Experimental Ophthalmology, 245(12): 1805–1814.

Hupkens, C., van den Berg, J., and van der Zee, J., 1999. National health interview surveys in Europe: an overview. Health Policy, 47**:** 145-168.

Huppert, F.A., Johnson, T., and Nickson, J., 2001. High prevalence of prospective memory impairment in the elderly and in early-stage dementia: findings from a population based study. In: L. Kvavilashvili and J. Ellis (eds), 'New Perspectives in Prospective Memory', Applied Cognitive Psychology, Special Issue, 14: 1–19.

Huppert, F.A., 2003. Designing for Older Users. In: J Clarkson, R Coleman, S Keates & C Lebbon (eds) Inclusive Design: Design for the whole population. Springer Verlag, London.

Huppert, F.A., Gardener, E., and McWilliams, B., 2006. Cognitive function. In: Banks et al (eds) Retirement, health and relationships of older population in England: The 2004 English Longitudinal Studies of Ageing (Wave 2). Institute for Fiscal Studies. London.

Johnson, D., Clarkson, J., and Huppert, F.A., 2009. Capability measurement for Inclusive Design, Journal of Engineering Design, First published on: 11 November 2009.

Kivinen, P., Sulkava, R., Halonen, P., and Nissinen, A., 1998. Self-reported and performance-based functional status and associated factors among elderly men: The Finnish cohorts of the seven countries study. Journal of Clinical Epidemiology, 51: 1243-1252.

Kempen, G.I., van Heuvelen, M.J., van den Brink, R.H., Kooijman, A.C., Klein, M., Houx, P.J., Ormel, J., 1996. Factors affecting contrasting results between self-reported and performance-based levels of physical limitations, Age and Ageing, 25(6): 458-64.

Laitinen, A., Koskinen, S., Harkanen, T., Reunanen, A., Laatikainen, L., Aromaa, A., 2005. Nationwide Population-Based Survey on Visual Acuity, Near Vision, and Self-Reported Visual Function in the Adult Population in Finland, Ophthalmology, 112(12): 2227-2237.

Latham, K.J.C., 1998. Who uses contrast sensitivity in optometric practice? Ophthalmic and Physiological Optics, 18: S2–13.

Lawton, M.P., and Brody, E.M., 1969. Assessment of older people: Self-maintaining and instrumental activities of daily living, Gerontologist, 9: 179-186.

Marmot, M., Banks, J., Blundell, R., Lessof, C., Nazroo, J., (eds.), 2003. Health, wealth and lifestyles of the older population in England: The 2002 English Longitudinal Study of Ageing, London: IFS.

McGinley, C., Bound, J., and Dong, H., 2010. Contextualising Data for Design. In: Proceedings of the 5th Cambridge Workshop on Universal Access and Assistive Technology (CWUAAT), Cambridge.

Molenbroek, J.F.M., 1987. Anthropometry of Elderly People in the Netherlands; Research and Applications. Applied Ergonomics, 18(3):187-199.

Myers, A.M., Holliday, P.J., Harvey, K.A.., Hutchinson, K.S., 1993. Functional performance measures: are they superior to self-assessments? Journal of Gerontology, 48: M196-206.

Norris, B., and Wilson, J., 1995. CHILDATA: The handbook of child measurements and capabilities - Data for design safety London: Department of Trade and IndustPeebles, L., and Norris, B., 1998.

ADULTDATA: The handbook of adult anthropometric and strength measurements - data for design safety. London: Department of Trade and Industry.

Pelli, D.G., Robson, J.G., and Wilkins, A.J., 1988. The design of a new letter chart for measuring contrast sensitivity, Clinical Vision Science, 2(3): 187–99.

Pennathur, A., Contreras, L. R., Arcaute, K., and Dowling, W., 2003. Manual dexterity of older Mexican American adults: A cross-sectional pilot investigation. International Journal of Industrial Ergonomics, 32(6): 419-431.

Porter, J.M., Case, K., Marshall, R., Gyi, D.E. and Sims, R.E., 2004. Beyond Jack and Jill: Designing for individuals using HADRIAN. International Journal of Industrial Ergonomics, 33: 249-264.

Rubin, G.S, Bandee-Roche, K., Huang, G.H., Munoz, B., Schein, O.D., Fried, L.P., 2001. The association of multiple visual impairments with self-reported visual disability: SEE project, Investigative Ophthalmology and Visual Science, 42: 64–72.

Salthouse, T.A., 1991. Theoretical perspectives on cognitive ageing. Lawrence Erlbaum Associates, Hillsdale, NJ.

Scheier, M.F., Carver, C.S., and Bridges, M.W., 1994. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test, Journal of Personality and Social Psychology, 67: 1063-1078.

Schwarzer, R., 1993. Measurement of perceived self-efficacy. Psychometric scales for cross-cultural research, Berlin: Freie Universität Berlin.

Smith, S., Norris, B., and Peebles, L., 2000. OLDER ADULTDATA: The handbook of measurements and capabilities of the older adult - Data for design safety London: Department of Trade and Industry.

Steel, N., Huppert, F. A., McWilliams, B., and Melzer, D., 2003a. Physical and cognitive function. In: Marmot, M., Banks, J., Blundell, R., Lessof, C., and Nazroo, J., (eds). Health, Wealth and Lifestyles of the Older Population in England: The 2002 English Longitudinal Study of Ageing, London: Institute for Fiscal Studies.

Steel, N., Maisey, S., and Cox, K., 2003b. Quality of healthcare. In: Marmot, M., Banks, J., Blundell, R., Lessof, C., and Nazroo, J., (eds). Health, wealth and lifestyles of older population in England: The 2002 English Longitudinal Studies of Ageing. Institute for Fiscal Studies. London.

Steenbekkers, L.P.A and van Beijsterveldt, C.E.M., 1998. Design-relevant Characteristics of Ageing users. The Netherlands: Delft University Press, Delft.

Tate, R., Smeeth, L., Evans, J., Fletcher, A., 2005. The prevalence of visual impairment in the UK. Report commissioned by the Royal National Institute of the Blind. http://www.rnib.org.uk

Thayaparan, K., Crossland, M.D., and Rubin, G.S., 2007. Clinical assessment of two new contrast sensitivity charts, The British Journal of Ophthalmology, 91: 749–752.

Thomas, P.D., Hunt, W.C., Garry, P.J., Hood, R.B., Goodwin, J.M., Goodwin, J., 1983. Hearing acuity in a healthy elderly population: effects on emotional cognitive and social status. Journal of Gerontology, 38: 321-5.

Thomson, D., 2005. VA testing in optometric practice. Part 2: Newer chart designs. Optometry Today: 22-24.

Tiffin, J., and Asher, E.J., 1948. The Perdue Pegboard: norms and studies of reliability and validity, Journal of Applied Psychology, 32: 232-247.

Voeks, S.K., Gallagher, C.M., Langer, E.H., Drinka, P.J., 1993. Self-reported hearing difficulty and audiometric thresholds in nursing home residents, Journal of Family Practice, 36(1): 54-8.

Weinstein, B.E., and Ventry, I.M.., 1982. Hearing impairment and social isolation in the elderly. Journal of speech and hearing research, 25: 593-9.

Table 1: Summary of different measures used in the study

	Self-report	Performance test
Component functions		
Vision	√	√
Hearing	√	√
Dexterity	√	√
Cognitive function	√	-
Component activities		
Reading	√	√
Speech recognition	√	√
Manual co-ordination	√	√
Psychological characteristics		
General	√	-
Product-related	√	-
Personality	√	-
Product interactions		
Iron	√	√
Microwave	√	√
Landline telephone	√	√

Table 2: An overview of the sample characteristics

	50-64 (N = 34)	65-80 (N=66)	All (N=100)
Gender			
Male	47.1	42.4	44.0
Female	52.9	57.6	56.0
Education			
Degree or Equivalent	47.1	53.0	51.0
Intermediate or No qualifications	52.9	47.0	49.0
Living arrangements			
Alone	26.5	28.8	28.0
With Others	73.5	71.2	72.0
Employment Status			
Currently employed	23.5	6.1	12.0
Not currently employed	76.5	93.9	88.0
Occupation			
Professional and Managerial	58.8	45.5	50.0
Intermediate	29.4	31.8	31.0
Manual and Routine	11.8	22.7	19.0
Household Income			
Less than £10,000	2.9	18.2	13.0
£10,000 - £50,000	91.2	69.7	77.0
Above £50,000	5.9	12.1	10.0
Ethnicity			
White	34.0	64.0	98.0
Non-white	0.0	2.0	2.0

Table 3: Summary statistics for component functions, by age

Component functions characteristics			***50-64*** *(N = 34)*	***65-80*** *(N=66)*	***All***
Vision	Self-report (% Fair or Poor)	General vision	2.9	9.1	7.0
		Distant vision	5.9	9.1	8.0
		Close vision	11.8	21.2	18.0
	Performance (Mean, SD)	Visual acuity *(Snellen, metres)*	6/6.6 (2.00)	6/7.5 (1.81)	6/7.2 (1.92)
		Contrast sensitivity *(log units)*	1.93 (0.45)	2.39 (0.68)	2.23 (0.65)
Hearing	Self-report (% Fair or Poor)	General hearing	2.9	27.3	19.0
		Hearing with background noise	32.1	43.9	40.0
		Distant hearing	2.9	41.2	15.0
	Performance (Mean, SD)	Hearing threshold *(dB)*	16.9 (6.10)	23.9 (12.38)	21.5 (11.14)
Dexterity	Self-report (% Fair or Poor)	Finger movements	5.8	12.1	10.0
		Picking-up objects	11.7	16.7	13.0
		Tying objects	11.8	10.6	11.0
		Turning objects	5.9	6.0	6.0
		Twisting objects	23.5	30.3	28.0
	Performance (Mean, SD)	Fine finger dexterity *(no. of parts)*	29.6 (6.09)	26.8 (6.03)	27.75 (6.16)
Cognitive function	Performance (Mean, SD)	Immediate memory recall	6.9 (1.33)	6.1 (1.63)	6.3 (1.58)
		Delayed memory recall	5.9 (1.61)	5.2 (1.95)	5.5 (1.87)
		Prospective memory *(% success)*	70.6	60.6	64.0
		Symbol learning	8.3 (1.69)	6.4 (2.44)	7.0 (2.38)
		Search speed *(no. of letters)*	309 (54.60)	295 (68.50)	300 (64.20)
		Search accuracy *(no. of letters missed)*	3.0 (2.38)	3.8 (2.63)	3.5 (2.56)
		Literacy	3.6 (0.65)	3.6 (0.68)	3.6 (0.67)

Table 4: Summary statistics for component activities, by age

	Component activities characteristics	50-64	65-80	All
Reading	1. Reading text on digital screen			
	Self-report (Mean, SD):			
	Frequency	7.2 (2.10)	6.8 (2.15)	6.9 (2.12)
	Difficulty	2.2 (1.39)	2.7 (1.56)	2.6 (1.51)
	Performance (% success):			
	High luminance	97.1	100.0	99.0
	Medium luminance	97.1	95.5	96.0
	Low luminance	20.6	10.6	14.0
	2. Reading text on plastic surface			
	Self-report (Mean, SD):			
	Frequency	6.1(1.87)	5.9 (1.52)	6.0 (1.64)
	Difficulty	3.3 (1.86)	3.9 (1.54)	3.7 (1.66)
	Performance (% success)	58.8	36.4	44.0
Speech recognition	3. Hearing messages at different volumes			
	Self-report (Mean, SD):			
	Frequency	5.3 (2.13)	4.9 (2.20)	5.1 (2.17)
	Difficulty hearing	2.9 (1.50)	3.2 (1.48)	3.1 (1.49)
	Performance (% success):			
	High volume	97.1	89.4	88.0
	Medium volume	94.1	84.8	88.0
	Low volume	91.2	74.2	80.0
Manual co-ordination	4. Twisting dials			
	Self-report (Mean, SD):			
	Frequency	6.1 (1.91)	6.1 (2.10)	6.1 (2.03)
	Difficulty	2.1 (1.41)	2.1 (1.30)	2.1 (1.30)
	Performance (% success):	64.7	45.5	52.0
	5. Using a touch screen key pad			
	Self-report (Mean, SD):			
	Frequency	3.7 (2.18)	2.6 (1.86)	3.0 (2.03)
	Difficulty	2.3 (1.30)	3.2 (1.58)	2.8 (1.53)
	Performance (% success)	88.2	68.2	75.0
	6. Pressing buttons on a remote control			
	Self-report (Mean, SD):			
	Frequency	7.1 (1.83)	7.3 (1.40)	7.2 (1.55)
	Difficulty	2.9 (1.55)	3.6 (1.60)	3.4 (1.61)
	Performance (% success)	76.5	48.5	58.0

Table 5: Summary statistics for psychological characteristics, by age

Psychological characteristics		***50-64*** *(N = 34)*		***65-80*** *(N=66)*		***All***	
		Mean	SD	Mean	SD	Mean	SD
General	General Self-efficacy	5.4	1.18	5.4	1.32	5.4	1.27
	General Self-esteem	5.0	1.18	4.8	1.26	4.8	1.23
	General Optimism	5.3	1.50	4.9	1.19	5.0	1.31
	General Mastery	4.8	1.86	5.3	1.80	5.1	1.83
Product-related	Product Self-efficacy	5.5	1.42	5.4	1.29	5.5	1.32
	Product Self-esteem	4.5	1.45	4.5	1.06	4.5	1.20
	Product Optimism	5.1	1.70	5.0	1.56	5.0	1.60
	Product Mastery	5.6	1.61	5.4	1.69	5.4	1.66
Personality	Openness to new experiences	4.8	1.03	5.3	1.10	5.1	1.09
	Extraversion	4.1	1.48	4.3	1.49	4.2	1.48
	Conscientiousness	6.0	1.00	5.7	0.95	5.8	0.97
	Agreeableness	5.2	1.23	5.1	1.13	5.2	1.16
	Emotional Stability	4.8	1.26	4.9	1.31	4.9	1.28

Table 6: Summary statistics for product interactions, by age

Product interaction characteristics			***50-64***	***65-80***	***All***
Iron	Self-report	Frequency (Mean, SD)	4.4 (1.71)	4.0 (1.74)	4.1 (1.73)
		Difficulty (Mean, SD)	2.0 (1.49)	2.1 (1.30)	2.1 (1.36)
	Performance	Task 1: Successful completion (%)	94.1	89.4	91.0
		Task 1 : Time taken in sec (Mean, SD)	23.6 (14.57)	26.2 (13.98)	25.3 (14.61)
		Task 2: Successful completion (%)	91.2	90.9	91.0
		Task 2 : Time taken in sec (Mean, SD)	13.6 (10.20)	15.5 (11.10)	14.8 (10.81)
Microwave	Self-report	Frequency (Mean, SD)	5.7 (2.15)	6.1 (2.22)	6.0 (2.20)
		Difficulty (Mean, SD)	2.1 (1.08)	2.2 (1.43)	2.1 (1.32)
	Performance	Task 1: Successful completion (%)	82.4	66.7	72.0
		Task 1 : Time taken in sec (Mean, SD)	56.3 (47.50)	52.1 (35.10)	53.5 (39.55)
		Task 2: Successful completion (%)	11.8	7.6	9.0
		Task 2 : Time taken in sec (Mean, SD)	83.6 (37.91)	83.4 (44.5)	83.4 (42.19)
Landline phone	Self-report	Frequency (Mean, SD)	7.4 (1.08)	7.2 (0.93)	7.3 (0.98)
		Difficulty (Mean, SD)	1.6 (1.23)	1.7 (1.11)	1.7 (1.15)
	Performance	Task 1: Successful completion (%)	20.6	12.1	15.0
		Task 1 : Time taken in sec (Mean, SD)	43.9 (33.10)	40.0 (32.04)	41.4 (32.30)
		Task 2: Successful completion (%)	64.7	50.0	55.0
		Task 2 : Time taken in sec (Mean, SD)	58.3 (23.51)	80.74 (47.75)	73.1 (42.38)

Table 7: Relationships among self-reported and performance measures

	Self-report Vs Performance test	*Spearman correlation*	*p-value*
Component functions	General vision Vs Visual acuity	**0.26**	0.009
	Distant vision Vs Visual acuity	**0.27**	0.006
	Close vision Vs Visual acuity	**0.29**	0.002
	General hearing Vs Hearing level	**0.50**	p< 0.001
	Hearing with noise Vs Hearing level	**0.30**	0.014
	Distant hearing Vs Hearing level	**0.32**	0.001
	Finger movements Vs Fine finger dexterity	**0.22**	0.029
	Picking objects Vs Fine finger dexterity	**0.40**	p<0.001
	Tying objects Vs Fine finger dexterity	**0.40**	p<0.001
	Turning objects Vs Fine finger dexterity	0.28	0.004
Component activities	Difficulty Vs Reading text on a digital screen	-0.12	0.28
	Difficulty Vs Using a touch screen keypad	**0.40**	0.001
	Difficulty Vs Hearing mobile messages	0.02	0.85
	Difficulty Vs Reading text on plastic surface	0.20	0.08
	Difficulty Vs Twisting a dial	**0.40**	p<0.001
	Difficulty Vs Pressing buttons on a remote control	**0.24**	0.019
Product interactions	Difficulty Vs Performance for Iron tasks	-0.10	0.40
	Difficulty Vs Performance for Microwave tasks	-0.10	0.40
	Difficulty Vs Performance for Landline tasks	-0.30	0.009

Table 8: Summary of significant variables from initial separate logistic regression analysis on each category of predictor variables for each of the three outcome measures*

	Product	Significant variables at 5% significance level	Odd's ratio (OR)	95% confidence intervals (CI)
Socio-demographics	Iron	Age	0.86	0.77, 0.96
		Gender	4.20	1.11, 15..8
	Microwave	Age	0.89	0.81, 0.98
		Household income	1.75	0.33, 9.30
	Landline phone	Age	0.93	0.86, 0.99
		Occupation	0.54	0.21, 1.41
Component functions	Iron	Self-report twisting objects	0.49	0.26, 0.94
		Self-report finger movements	4.10	1.38, 12.27
		Visual acuity	31.05	1.03, 931.26
		Hearing threshold	0.91	0.86, 0.96
		Visual search speed	1.01	1.00, 1.02
		Prospective memory	1.78	0.03, 0.97
	Microwave	Contrast sensitivity	409.83	6.11, 27479.48
	Landline phone	Self-report General vision	0.48	0.27, 0.88
		Self-report Close vision	2.76	1.25, 6.11
		Visual acuity	28.77	2.99, 276.93
		Symbol recall	1.19	0.99, 1.44
		Memory recall	0.82	0.69, 0.99
		Prospective memory	0.37	0.15, 0.94
Component activities	Iron	Self-report Frequency of reading text on digital screen	0.62	0.42, 0.91
		Self-report Difficulty in twisting dials	1.90	0.93, 4.04
		Successful task completion of pressing buttons on remote control	0.31	0.09, 1.02
		Successful task completion of hearing at low volume	22.0	1.68, 287.19
	Microwave	Successful task completion of using touch screen	0.34	0.12, 0.99
	Landline phone	Self-report Frequency of reading text on digital screen	1.70	1.06, 2.69
		Self-report Difficulty of reading text on digital screen	0.68	0.47, 0.98
		Successful task completion of hearing at low volume	4.20	1.4, 14.00
Psychological characteristics	Iron	Product Self-efficacy	1.79	1.64, 2.80
		General optimism	0.56	0.31, 0.98
		Openness to new experiences	0.49	0.25, 0.94
	Microwave	General Self-efficacy	0.65	0.41, 1.01
	Landline phone	General Mastery	1.34	1.09, 1.73
		General Self-efficacy	0.50	0.24, 1.02
		Product Self-efficacy	1.74	1.2, 2.5
		Openness to new experiences	0.41	0.21, 0.8

* The outcome measures:

1. at least one phone task completed
2. both iron tasks completed
3. at least one microwave task completed and

Table 9: Summary of significant predictors of product interaction from final logistic regression analysis. The variables entered in this analysis were the significant predictors from each category of predictor variables for each of the three outcome measures (reported in Table 8)

Product	***Significant variables at 1% significance level***	***Odd's ratio (OR)***	***99% confidence intervals (CI)***
Iron	Self-report Frequency of reading text on a digital screen	0.69	0.50, 0.94
	Hearing threshold	0.92	0.85, 1.00
Microwave	Contrast sensitivity	409.83	1.63, 103016
Landline telephone	Visual acuity	22.31	1.05, 475.73
	Openness to new experiences	2.44	1.02, 5.85
	Self-report Frequency of reading text on digital screen	0.62	0.42 , 0.93
	Successful task completion of hearing at low volume	5.19	1.12 , 23.96